Realizing Unbiased Deformation: A Theoretical Consideration

Alex Leow, Ming-Chang Chiang, S. C. Huang, Arthur W. Toga, Paul M. Thompson

Realizing Unbiased Deformation: A Theoretical Consideration

A.D. Leow, M.C. Chiang, S.C. Huang, A.W. Toga, and P.M. Thompson

***Abstract*— Maps of local tissue compression or expansion are often recovered by comparing MRI scans using nonlinear registration techniques. The resulting changes can be analyzed using tensor-based morphometry (TBM) to make inferences about anatomical differences. Numerous deformation techniques have been developed, although there has not been much theoretical development examining the mathematical/statistical validity of each technique. In this paper, we propose a basic principle that any registration technique should satisfy: realizing unbiased test statistics under null distribution of the displacement. In other words, any registration technique should recover zero change in the test statistic when comparing two images differing only in noise. Based on this principle, we propose a fundamental framework for the construction and analysis of image deformation. Moreover, we argue that logarithmic transform is instrumental in the analysis of deformation maps. Combined with the proposed framework, this leads to a theoretical connection between image registration and other branches of applied mathematics including information theory and grid generation.**

***Index Terms*-Mutual information, Image registration, Computational anatomy.**

1. Introduction

Non-linear image registration is a well-established field in medical imaging with many applications in functional and anatomic brain mapping, image-guided surgery, and multimodality image fusion [1-8]. The goal of image registration is to align, or spatially normalize, one image to another. In multi-subject studies, this serves to reduce subject-specific anatomic differences by deforming individual images onto a population average brain template.

The deformations that map each anatomy onto a common template can be analyzed voxel-wise to make inferences about relative volume differences between the individuals and the template or statistical differences in anatomy between populations [9-13]. Similarly, in longitudinal studies it is possible to visualize structural brain changes that occur over time by deforming subjects' baseline scans onto their subsequent scans, using the deformation map to quantify local changes. This area of computational anatomy is known as *tensor-based morphometry* [11-15]. As there are numerous techniques for non-linear registration, one may ask whether some approaches are better than others in practice and/or in theory. In the following section, we will try to answer this fundamental question from a statistical standpoint.

2. Criteria for proper image registration

One could not study non-linear registration without closely examining the common practice of applying logarithmic transformation to Jacobian maps. Log-transformation of a Jacobian determinant field has become standard practice in most tensor-based morphometry (TBM) papers [11,12]. The Jacobian determinant of a diffeomorphic (smooth) map is bounded below by zero but unbounded above. Thus, at any voxel, its null distribution would be a better fit to a symmetric Normal distribution if the Jacobians are logged. Additional arguments that favor log-transformation of Jacobian maps come from the need to symmetrize the probabilities of expansions and shrinkages that are reciprocals of each other. We will discuss two different, yet related, concepts of symmetrizing; each supports the use of log transformation.

2.1. Two types of symmetry favor the log transformation of Jacobian maps

The first symmetry operates on a voxel level. The motivation behind this approach is the result of voxel-wise statistical tests. Consider, for example, testing for the presence of voxel-wise mean structural change in a multi-subject experiment. One might want to employ classical statistical approaches and test the null hypothesis of zero mean change using one-sample Student's *t* tests. A more general linear model typically relies on the residuals being a good approximation to a Gaussian random field. This null model assumes that at each voxel the observed rates of change over time (or relative volumes in a cross-sectional study) collected from all subjects follow a normal distribution, thus suggesting the use of log-transform – otherwise the Jacobian determinants are bounded below by zero but unbounded above.

The second type of symmetry is the symmetry of Jacobian distributions inside a region (in contrast, the symmetry previously described deals with multiple observations across subjects, at a single voxel). This second, albeit less intuitive, level of symmetry sheds light on how one might construct better non-linear image registrations.

In theory, a proper registration algorithm should produce unbiased estimators of the real anatomical change. An unbiased algorithm should detect no statistically significant change between two serial images if there is no difference other than noise, and statistics quantifying change collected from all voxels should ideally have a zero mean value; rates of change deviating from zero should be considered as errors. In a classical statistical setting, one might expect that rates of change thus follow a Gaussian distribution with zero mean, again justifying the use of the logarithmic transform.

These two types of symmetry operate at different levels: one across all voxels when comparing two images, and the other across all subjects at one single voxel. Assuming this symmetric log-normal distribution for the residuals in a statistical model for the Jacobian determinant, one symmetrizes the rate of change distribution by considering a halving or doubling of volume to be equally likely *a priori.* A related approach is taken by Pennec et al. [16], where the Cauchy-Green strain tensor [17, 18] of a deformation mapping is logged and used as a term in a penalty functional that is integrated over the image domain to regularize the deformation.

2.2. Detecting no change in the absence of real change as a necessary criterion for proper registrations: realizing unbiased test statistics

As noted previously, the logarithmic transform is fundamental in analyzing deformation maps through Jacobian maps, and an ideal registration algorithm should yield Jacobian maps and test statistics that imply zero-change, when no difference is present between two registered images. We refer to this as the *principle of realizing unbiased test statistics under null distribution of the displacement.*

Before we show how to construct unbiased test statistics, we first define both T and S, on an image domain Ω, as the two images to be registered. Let us also assume, without loss of generality, that the volume of this domain is 1, i.e., $|\Omega|=1$. We seek to estimate a transformation h such that S is non-linearly registered to T when deformed by h (i.e., $S(h(x))$). In this paper, we will restrict this mapping to be differentiable, one-to-one, and onto from the image domain onto itself [19] (in practice/implementation, the one-to-one and onto property can be approximated by extending the boundary towards infinity). Let us associate three probability density functions, defined on this image domain, to the identity mapping (id) as well as the deformation h and its inverse:

$$pdf_h(x) = |Dh(x)|;\ pdf_{h^{-1}}(x) = |Dh^{-1}(x)|;\ pdf_{id}(x) = 1. \tag{1}$$

Here, the Jacobian matrix of a transformation h is denoted Dh. As noted in [19], integrating the log of the Jacobian determinant over the image domain simply calculates the Kullback-Leibler distance between the identity map and h.

$$\int_\Omega \log|Dh(x)|dx = -KL(pdf_{id}, pdf_h) \le 0. \tag{2}$$

Here KL, the non-negative asymmetric Kullback-Leibler (K-L) distance, between two PDFs p and q, is defined as $KL(p,q) = \int_\Omega p \log \frac{p}{q} dx \ge 0;$

Given Eq. (2), one might ask if we could use the negative integral of the logged Jacobian determinant $-\int_\Omega \log|Dh(x)|dx$ as a basis for regularization. This integral only evaluates to zero when h is volume-preserving everywhere. Minimizing this integral gives us a volume-preserving map, and thus realizes unbiased statistics. On a global scale, this integral also evaluates the overall mean log-Jacobian with respect to the computation domain. Moreover, in the field of information theory, the KL distance is geometrically important, providing a means to compare probability density functions on general manifolds [20].
Interestingly, the integral Eq.(2) has skew-symmetry with respect to h and its inverse

$$\begin{aligned} KL(pdf_{id}, pdf_h) &= \int -\log|Dh|dx \\ = KL(pdf_{h^{-1}}, pdf_{id}) &= \int |Dh^{-1}|\log|Dh^{-1}|dx \end{aligned} \tag{3}$$

Notably, the idea of integrating the square root of the Jacobian, as proposed in [21], to remove the skew-symmetry, is equivalent to calculating the Bhattacharyya distance ***B,*** a well-known measure in information theory [21,22].

$$B(pdf_{id}, pdf_h) = \int \left|Dh(x)\right|^{1/2} dx$$
$$= \int \left|Dh^{-1}(x)\right|^{1/2} dx = B(pdf_{h^{-1}}, pdf_{id}) \quad \textbf{(4)}$$

Surprisingly, notice that here the Bhattacharyya distance is symmetrical with respect to its two arguments, as well as inverse-consistent. To connect the KL-distance and Bhattacharyya distance, one can consider the geodesic linking of the two PDFs: $P(.,t)$, parameterized by time t:

$$P(x,t) = \frac{pdf_{id}(x)^t \, pdf_h(x)^{1-t}}{N}, \quad N = \int pdf_{id}{}^t(x)\, pdf_h(x)^{1-t}\, dx \quad \textbf{(5)}$$

The Bhattacharyya distance corresponds to the arbitrary choice of t=1/2, while a generalization of the above leads to the Chernoff distance in information theory [23].

2.3. Realizing unbiased deformation using symmetrical KL distance

To construct unbiased deformations, we first generalize eq. (2): given any diffeomorphism g mapping Region of interests S (with size a) to T (with size b), we have the following $\int_S \log|Dg(x)|\,dx \le \log b/a$, while equality holds if and only if the Jacobian of g takes the constant value b/a. The implication of this generalization is that, assuming ROI mapping from S to T, the Log operation is unbiased if and only if the corresponding Jacobian field is evenly distributed. Thus, minimizing the negative integral of logged Jacobian (treated as a cost function) as in 2.2 again leads to unbiased statistics. Given this generalization, let us now introduce the following by combining **Eq.(3)**, and its counterpart contributed by the inverse mapping:

$$KL(pdf_h, pdf_{id}) + KL(pdf_{h^{-1}}, pdf_{id}) = KL(pdf_h, pdf_{id}) + KL(pdf_{id}, pdf_h)$$
$$= KL(pdf_{id}, pdf_{h^{-1}}) + KL(pdf_{id}, pdf_h) = KL(pdf_{id}, pdf_{h^{-1}}) + KL(pdf_{h^{-1}}, pdf_{id}) \quad \textbf{(6)}$$
$$= \int \left(\left|Dh(x)\right| - 1\right)\log\left|Dh(x)\right| dx = \int \left(\left|Dh^{-1}(x)\right| - 1\right)\log\left|Dh^{-1}(x)\right| dx$$

In this symmetric form, the integrand $\left(\left|Dh\right| - 1\right)\log\left|Dh\right|$ is always non-negative, compared to the integrand in **Eq. (3)** where locally negative numbers can be obtained (when the Jacobian is greater than zero), though globally non-negative. The unbiased property has two layers of meaning: the first one being realizing null distribution when comparing identical images differing in noise (from **Eq**. (2)); the second being realizing correct log Jacobian statistics in ROI analysis (from the generalization of **Eq**. (2)).

Under this framework, constructing deformations can be viewed as quantifying the symmetric KL distance between the identity map and the resulting deformation (or the inverse deformation due to its symmetry). Moreover, this framework embeds statistical analyses into the construction of deformations, penalizing deformations that skew the distribution of test statistics. A second interpretation of **Eq.(6)** is that it simply calculates the mean log Jacobian for h and its inverse inside the domain, thus encoding regional volume changes. Let us also point out an interesting observation by applying the square root Jacobian integrand as in **Eq.(4)** to the log-transformed Jacobian framework. A simple change of variables verifies the following:

$$\int \left(\log\left|Dh\right|\right)\left|Dh\right|^{1/2} dx + \int \left(\log\left|Dh^{-1}\right|\right)\left|Dh^{-1}\right|^{1/2} dx = 0 \qquad \textbf{(7)}$$

To further link **Eq.(6)** to other branches of mathematics, optimization problems involving Jacobian operator are commonly encountered in grid generation [24] and in continuum mechanics, where the Hencky tensor arises as logged tensor parameters in modeling very large deformations. However, we believe that the logarithmic transform has not been formally introduced in the grid generation literature and may also be useful there.

2.4. Logarithmic transform: a ubiquitous operation

With **Eq. (6),** realizing unbiased log-Jacobian values can be thought of as equivalent to minimizing the KL-distance. Moreover, in this case the concept of inverse-consistency translates to the symmetrization of the KL-divergence. The choice of the logarithmic operation in defining KL-divergence, though arbitrary at first sight, can now be easily justified by linking it to the statistical analyses of deformations in non-linear registration.

Finally, we comment briefly on how this framework can be implemented. Given an image matching cost function *C*, we seek, among all deformations minimizing this matching cost, the deformation with minimal symmetric KL distance as defined in Eq (6). In practice, this often means implementing a Lagrange multiplier, resulting in the following combined minimization problem

$$\underset{h\in H}{\arg\min} C(T, T\circ h^{-1}, S, S\circ h) + \lambda\left(KL(pdf_h, pdf_{id}) + KL(pdf_{id}, pdf_h)\right) \qquad \textbf{(8)}$$

Here λ is the multiplier, and *H* is the solution space, of which a common choice consists of all one-to-one, onto, and differentiable maps. Often, the solution is numerically obtained by recursive smoothing or regularization applied to the force field. In this paper, we focus on constructing a general principle that applies to different numerical approaches, and refrain from touching on the issue of regularization. The terms in **Eq.(6)** may also be viewed as regularizers, or deformation priors, as they penalize log-Jacobian values that deviate from zero. However, it is well-known in the grid generation field that an integral constraint on the Jacobian alone does not generally guarantee a globally smooth grid [24], so the smoothness of the resulting maps deserves further study.

Lastly, to complete our discussions on implementation issues, let us provide the gradient descent direction contributed by the symmetric KL distance term in eq(8), via its Euler-Lagrange equation. To this end, let us denote Co_{ij}, the matrix cofactor for the (i,j) th component of the Jacobian matrix Dh, we then obtain its Euler-Lagrange equation, using standard Calculus of variations with respect to the i-th coordinate as follows

$$\sum_j \frac{\partial}{\partial x_j}\left\{\left(1+\log|Dh(x)| - 1/|Dh|(x)\right)Co_{ij}(x)\right\} = 0, \qquad \textbf{(9)}$$

$$\left(Dh(x)\right)^{-1} = \left(Co_{ij}(x)\right)^T / |Dh(x)|.$$

Figure 1 shows a numerical example demonstrating the proposed approach implemented using a gradient descent projection method. Here, we matched the sequential MRI images (both of size 128 by 128 by 128) from a single subject diagnosed with semantic dementia. The figure showed three 2D slices, plotting the source, target, as well as the deformed source images. As in other TBM approaches, the technique is valuable as the pattern of atrophy is computed automatically, without interactive specification of regions of interest.

3. Conclusion

This paper is the first effort to systemically examine the relationship of image registration, information theory, and grid generation. While information-theoretic measures such as mutual information [25-28], *f*-divergence [29] and Jensen-Rényi divergence [13] have been popular measures to describe intensity correspondences in nonlinear image registration, it is much less common to appeal to statistical divergence measures in analyzing deformation fields. The proposed formulation calculates the KL-divergence between the deformation and the identity map, treating them as density functions defined on images. Unlike approaches employing conventional continuum mechanics (e.g., Lamé coefficients), our formulation is unbiased and parameter-free (the only parameter involved is the weight, which may be viewed as a Lagrange multiplier). The approach does not therefore make strong assumptions to model the deformation process. In contrast, the commonly employed elasticity theory assumes that displacements or velocity fields to conform to the law of elasticity, which most likely does not accurately describe any real brain deformation process over time.

The proposed framework helps explain the need for log-transformation of Jacobian values in TBM studies, which is ubiquitous and essential in analyzing tissue shrinkage/expansion. Symmetrization is also fundamental to securing inverse-consistency, an important property in image registration. Lastly, this framework is also consistent with a large-deformation approach [4,6,7], as any one-to-one and onto diffeomorphism remains in the solution space (since its Jacobian is positive and finite everywhere, and thus is theoretically attainable). In the future, we intend to further investigate this approach, which we believe provides a new perspective in non-linear registration and connects it to other fields of mathematics.

References

[1] P. M. Thompson and A. W. Toga, "A framework for computational anatomy," *Computing and Visualization in Science*, vol. 5, pp. 13-34, 2002.

[2] U. Grenander and M. I. Miller, "Computational anatomy: An emerging discipline," *Quarterly of Applied Mathematics*, vol. 56, pp. 617-694, 1998.

[3] M. K. Chung, K. J. Worsley, T. Paus, C. Cherif, D. L. Collins, J. N. Giedd, J. L. Rapoport, and A. C. Evans, “A unified statistical approach to deformation-based morphometry. *NeuroImage*, vol. 14, pp. 595–606, 2001.

[4] B. Avants and J. C. Gee, “Geodesic estimation for large deformation anatomical shape averaging and interpolation,” *NeuroImage*, vol. 23, suppl. 1, S139-50, 2004.

[5] D. Shen, and C. Davatzikos, “Very high-resolution morphometry using mass-preserving deformations and HAMMER elastic registration,” *NeuroImage*, vol. 18, no. 1, pp. 28-41, 2003.

[6] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435-1447, 1996.

[7] M. I. Miller, "Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms." *NeuroImage*, vol. 23, suppl. 1, pp. S19-S33, 2004.

[8] D. L. Collins, T. M. Peters, and A. C. Evans, "Automated 3D nonlinear deformation procedure for determination of gross morphometric variability in human brain," *Proc. SPIE* 2359, pp. 180-190, 1994.

[9] C. Studholme, V. Cardenas, N. Schuff, H. Rosen, B. Miller, and M. Weiner, "Detecting spatially consistent structural differences in Alzheimer's and fronto-temporal dementia using deformation morphometry," *Proc. MICCAI*, pp. 41-48, 2001.

[10] P.M. Thompson et al. (2004). *Mapping Cortical Change in Alzheimer's Disease, Brain Development, and Schizophrenia*, NeuroImage, 23 Suppl 1:S2-18, September 2004.

[11] P. M. Thompson, J. N. Giedd, R. P. Woods, D. MacDonald, A. C. Evans, and A. W. Toga, "Growth patterns in the developing brain detected by using continuum mechanical tensor maps," *Nature*, vol. 404, no. 6774, pp. 190-3, 2000.

[12] P. Cachier and D. Rey. "Symmetrization of the Non-Rigid Registration Problem using Inversion-Invariant Energies: Application to Multiple Sclerosis." *In MICCAI'00, LNCS* 1935:472-481, Pittsburgh, PA, Oct. 2000.

[13] A. D. Leow, J. C. Soares, K. M. Hayashi, A. D. Klunder, A. D. Lee, C. E. Bearden, E. S. Monkul, M. A. Nicoletti, A. P. Cerchiari, M. Trakhenbroit, P. Brambilla, R. B. Sassi, A. G. Mallinger , A. W. Toga, and P. M. Thompson, (2005). *Asymmetrical Effects of Lithium on Brain Structure Mapped in Healthy Individuals*, submitted for publication.

[14] A. D. Leow, A. D. Klunder, C.R. Jack, A.W. Toga, A.M. Dale, M.A. Bernstein, P.J. Britson, J.L. Gunter, C.P. Ward, J.L. Whitwell, B. Borowski, A. Fleisher, N.C. Fox, D. Harvey, J. Kornak, N. Schuff, C. Studholme, G.E. Alexander, M.W. Weiner, P.M. Thompson, For the ADNI Preparatory Phase Study (2005). "Longitudinal Stability of MRI for Mapping Brain Change using Tensor-Based Morphometry," *NeuroImage*, in press.

[15] J. Ashburner, J. Anderson, and K. Friston, "High-dimensional image registration using symmetric priors," *NeuroImage*, vol. 9, pp. 619-628, 1999.

[16] X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache, "Riemannian elasticity: a statistical regularization framework for non-linear registration", *Proc. MICCAI 2005*, LNCS 3750, pp. 943-950, 2005.

[17] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Fast and simple calculus on tensors in the Log-Euclidean framework," Proc. *MICCAI 2005*, Palm Springs, CA, October 26-29, 2005.

[18] R. P. Woods, "Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation," *NeuroImage*, vol. 18 no. 3, pp. 769-88, 2003.

[19] Leow AD, Huang SC, Geng A, Becker JT, Davis SW, Toga AW, Thompson PM (2005). *Inverse Consistent Mapping in 3D Deformable Image Registration: Its Construction and Statistical Properties*, IPMI2005, Glenwood Springs, Colorado, July 11-15, 2005.

[20] N.N. Cencov. *Statistical Decision Rules and Optimal Inference*, Volume 14 of *Translations in Mathematics*. American Mathematical Society, Providence, RI, 1982

[21] M. Nielsen, P. Johansen, A Jackson, B Lautrup (2001). Statistical warps, a least committed model, MICCAI2001.

[22] A. Bhattacharyya. On a Measure of Divergence Between Two Statistical Populations Defined by their Probability Distributions. Bull. Calcutta Math. Soc., 35, 99 (1943)

[23] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[24] V.D. Liseikin, Grid Generation Methods. Springer-Verlag, Heidelberg, 1999.
[25] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, “A viscous fluid model for multimodal non-rigid image registration using mutual information,” *Medical Image Analysis*, vol. 7, pp. 565-575, 2003.
[26] B. Kim, J. L. Boes, K. A. Frey, and C. R. Meyer, “Mutual information for automated unwarping of rat brain autoradiographs,” *NeuroImage*, vol 5, no. 1, pp. 31-40, 1997.
[27] P. Lorenzen, B. Davis, and S. Joshi, “Model based symmetric information theoretic large deformation multi-modal image registration,” *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 720-723, 2004.
[28] C. J. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, and C. J. Taylor, “A Unified Information-Theoretic Approach to Groupwise Non-Rigid Registration and Model Building,” *IPMI 2005.*
[29] J. P. W. Pluim, J. B. Antoine Maintz, and M. A. Viergever, “*f*-information measures in medical image registration,” *IEEE Transactions on Medical Imaging* 23(12):1508-1516, 2004.

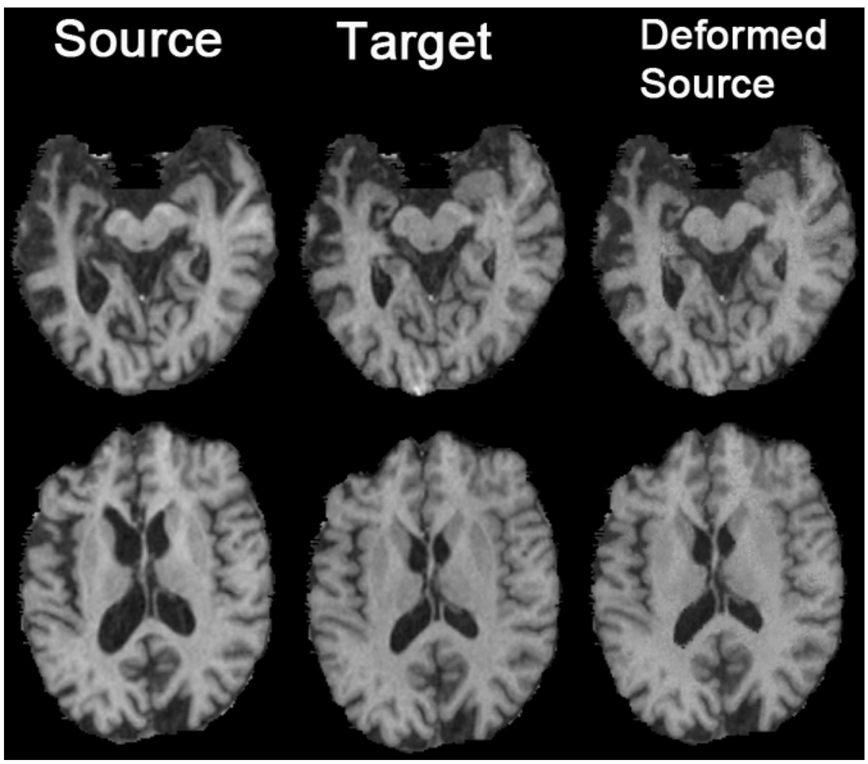

Fig. 1. Image registration using the unbiased deformation algorithm, with the sum of the squared intensity difference (SSD) as the cost function. 3D scans from a semantic dementia patient imaged at two time points were nonlinearly registered to estimate the profile of volumetric change. The later scan (first column) was chosen to be the source image. Progressive brain atrophy is observed in these axial slices.